

Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Konsumsi Produk Kosmetik milik PT Cedefindo

Amanda

Komputerisasi Akuntansi, STMIK Pranata Indonesia, amanda.saad1123@gmail.com

Moniq Veronica Sitorus

Teknik Informatika, STMIK Pranata Indonesia

Abstract

The increasing number of cosmetic product orders every year causes a lot of product data to be processed, causing difficulties in grouping the data. In this study applying Data Mining using the Clustering method to classify products that are often and which are rarely produced at PT. Cedefindo. The algorithm used is K-Means Clustering, where data grouped based on the same characteristics will be included in the same group and the data sets entered into the groups do not overlap. The information displayed is in the form of groups of product names and production quantities from September and October as a sample. The results of this study will assist the Company in analyzing which products are often produced and which are rarely produced. The software used to help this grouping is Rapid Miner.

Keywords: *K-Means, Clustering, Rapid Miner, Product*

Abstrak

Semakin meningkatnya jumlah order produk kosmetik setiap tahunnya menyebabkan banyaknya data produk yang perlu diolah sehingga menyebabkan kesulitan dalam pengelompokan data tersebut. Pada penelitian ini menerapkan Data Mining dengan menggunakan metode Clustering untuk mengelompokkan Produk yang sering dan yang jarang diproduksi di PT.Cedefindo. Algoritma yang digunakan yaitu K-Means Clustering, dimana data dikelompokkan berdasarkan karakteristik yang sama akan dimasukkan ke dalam kelompok yang sama dan set data yang dimasukkan ke dalam kelompok tidak tumpang tindih. Informasi yang ditampilkan berupa kelompok – kelompok nama produk dan jumlah produksi dari bulan September dan oktober sebagai sample. Hasil penelitian ini akan membantu pihak Perusahaan dalam menganalisa produk mana yang sering dan yang jarang diproduksi. Software yang digunakan untuk membantu pengelompokan ini adalah Rapid Miner.

Kata Kunci: *K-Means, Clustering, Rapid Miner, Produk*

I. PENDAHULUAN

Produk Kosmetik adalah bahan yang dipakai pada bagian luar tubuh (kulit, rambut, dan bibir) atau gigi dengan tujuan untuk membersihkan, mengharumkan, serta memperbaiki penampilan. Produk kosmetik salah satu produk yang sangat di gandrungi banyak orang pada zaman sekarang baik dari kalangan anak muda hingga dewasa.

Banyaknya penggunaan produk kosmetika menghasilkan banyaknya juga permintaan produksi pembuatan kosmetik sehingga menghasilkan banyaknya PO (*Pre-Order*) ke PT.Cedefindo, yang membuat banyaknya juga data produk yang terinput dan tak sering membuat data yang tak terkontrol.

Dalam penghitungan bahan kemas yang manual terkadang menghasilkan terjadinya penumpukan stok

barang kemas yang berlebih, terjadinya salah perkiraan antara jumlah PO (*purchase order*) *customer* dengan penyediaan bahan baku dan bahan kemas menghasilkan terjadinya selisih jumlah PO dengan bahan baku dan kemas sehingga hanya menghasilkan kerugian. Tidak tertatanya data produk yang diproduksi setiap bulannya. Belum lagi terjadinya kesalahan manusiawi dari karyawan staff, *quality control*, gudang, maupun produksi.

Metode Data Mining, K-Means, dan Clustering akan dipakai dalam memecahkan masalah penumpukan bahan baku dan kemas dan untuk mengelompokkan produk yang jarang dan sering diproduksi.

Data Mining merupakan bidang ilmu yang digunakan untuk menangani masalah pengambilan

informasi dari database yang besar dengan menggabungkan teknik dari statistik, pembelajaran mesin, visualisasi data, pengenalan pola, dan database. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistik data.

K-Means adalah algoritma clustering partisi sederhana berbasis prototipe yang mencoba menemukan cluster yang tidak tumpang tindih. Cluster ini diwakili dengan sentroid mereka (sentroid cluster biasanya merupakan rata-rata titik dalam cluster itu). Pada algoritma K-Means komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui dahulu target kelasnya.

Pengelompokan (*Clustering*) adalah suatu metode data mining yang berfungsi untuk mengelompokkan beberapa data sehingga menghasilkan satu data yang mirip hingga membentuk satu data yang baru.

Pengelompokan (*Clustering*) Produk Kosmetik yang sering dan jarang di produksi perlu dilakukan di PT.Cedefindo untuk memudahkan dalam merinci setiap Produk yang di produksi di perusahaan itu sendiri agar perusahaan dapat lebih meyiapkan jam kerja dan strategi kerja yang lebih efisien dalam melakukan setiap produksi produk.

Metode Algoritma K-Means Clustering dipakai karna metode pengelompokan data yang sederhana dan banyak digunakan dalam aplikasi data mining, (Zhang & Fang, 2013) melakukan penelitian dalam perbaikan algoritma K-means untuk mengoptimalkan inisialisasi pusat cluster. Dengan menemukan satu set data yang mencerminkan karakteristik distribusi data sebagai pusat awal cluster untuk mendukung pembagian data ke batas yang terbaik.

Namun dalam penghitungan bahan kemas yang manual terkadang menghasilkan terjadinya penumpukan stok barang kemas yang berlebih, terjadinya salah perkirain antara jumlah PO (*purchase order*) customer dengan penyediaan bahan baku dan bahan kemas menghasilkan terjadinya selisih jumlah PO dengan bahan baku dan kemas sehingga hanya menghasilkan kerugian. Tidak tertatanya data produk yang diproduksi setiap bulannya. Belum lagi terjadinya kesalahan manusiawi dari karyawan staff, *quality control*, gudang, maupun produksi.

Perusahaan manufaktur mendapatkan keuntungan dari setiap *man hours* (jam kerja) yang sudah di sepakati dari pihak *customer* dan perusahaan. Lalu untuk menentukan jam kerja yang efisien dibutuhkannya menentukan mana produk yang sering diproduksi dan mana yang jarang diproduksi agar setiap diproduksinya suatu produk kita bisa menentukan jam kerja yang pas untuk mengerjakan produk itu.

Dalam kesalahan perhitungan perkiraan itu maka bisa diatasi dengan beberapa metode salah satunya metode K-Means Clustering. Metode ini mengelompokkan setiap produk mana yang sering diproduksi dan mana yang jarang diproduksi. Dengan metode ini maka bisa meminimalisir penyetokan barang yang berlebih. Dimulai dari perencanaan waktu produksi, penyediaan bahan baku, dan bahan kemas. Dengan dikelompokkannya data produk mana saja yang sering dan yang jarang diproduksi membuat perkiraan jam kerja hingga kesalahan-kesalahan kecil dalam tahapan produksi.

II.METODE PENELITIAN

Teknik pengumpulan data yang dibutuhkan dalam penelitian yaitu data set produksi produk kosmetika PT.Cedefindo dari bulan September-Oktober 2020. Dengan menggunakan beberapa metode yaitu:

a) Pengamatan (Observasi)

Pengamatan (observasi) yang dimaksud dalam penelitian ini adalah melakukan penelitian secara langsung dengan tujuan untuk lebih memahami dan mengetahui langkah-langkah apa saja yang harus diambil dalam menyelesaikan permasalahan yang ditemukan.

b) Studi Pustaka

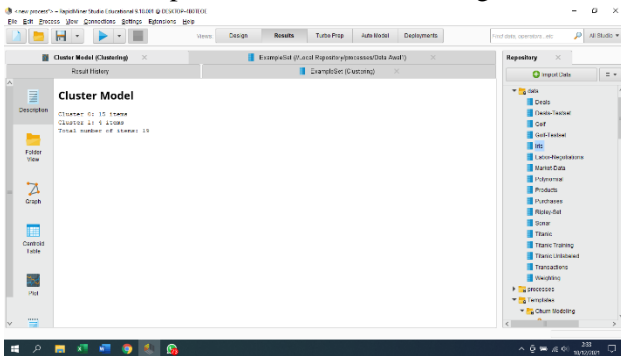
Pada penelitian juga menggunakan teknik pengumpulan data studi pustaka dilakukan dengan mempelajari teori-teori atau literatur dari buku-buku, jurnal, referensi-referensi yang berhubungan dengan K-Means untuk melengkapi data.

c) Memperoleh data melalui pengkajian dan penelaahan catatan-catatan maupun dokumen-dokumen yang berkaitan dengan masalah yang sedang diteliti.

III. HASIL DAN PEMBAHASAN

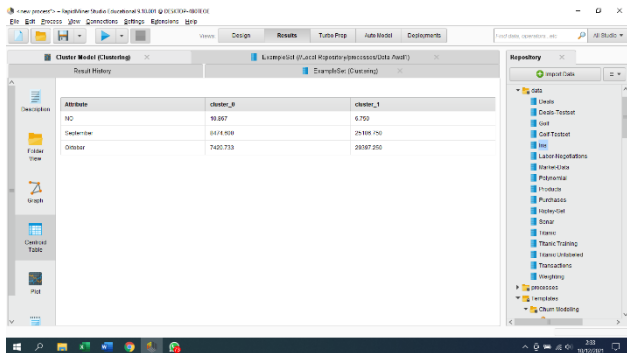
1. Proses K-Means Clustering

Berikut proses K-Means Clustering.



Gambar 1. Hasil Proses K-Means Clustering

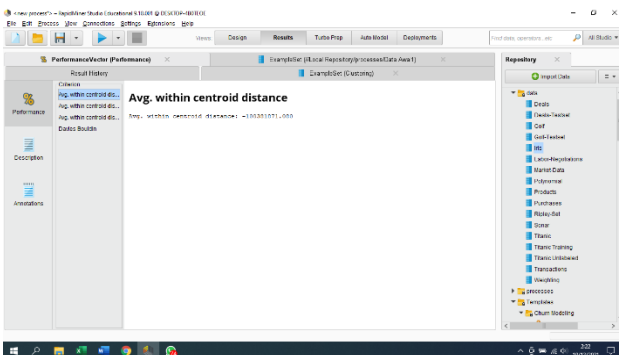
Hasil dari proses K-Means Clustering di dapat cluster model : Cluster 0 ; 15 items, Cluster 1 : 4 item dengan total 19 items. Lalu di dapat *Centroid table* :



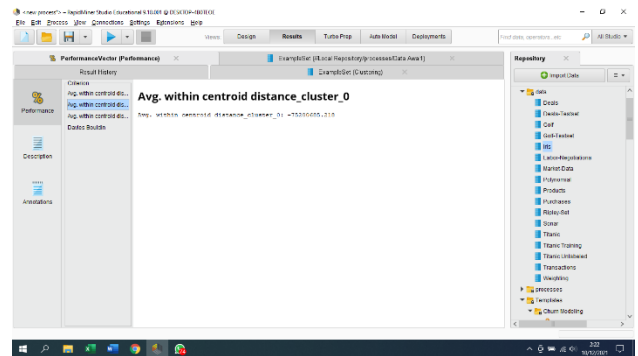
Gambar 2. Centroid Table

2. Proses Performance

Beberapa hasil proses *performance*



Gambar 3. Hasil Proses Avg. within centroid distance

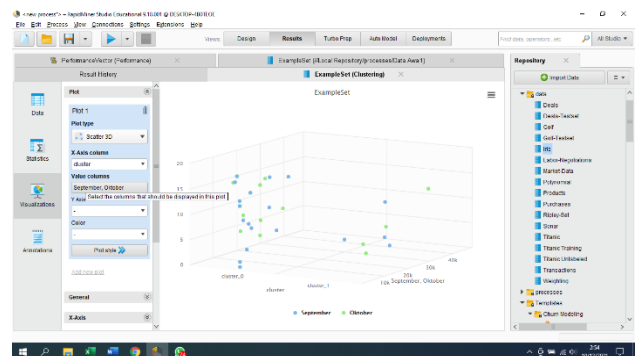


Gambar 4. Hasil Proses Avg. within centroid distance_cluster_0



Gambar 5. Hasil Proses Avg. within centroid distance_cluster_1

Dari hasil proses *Performance* maka di dapat *Example clustering* dalam bentuk *scatter 3D* berikut:



Gambar 6. Tampilan Scatter 3D

IV. KESIMPULAN

Berdasarkan hasil yang didapat dari proses dalam menentukan jumlah cluster terbaik dengan metode K-Means maka dapat disimpulkan sebagai berikut :

1. Pengklasteran dengan menggunakan metode K-Means didapatkan untuk menentukan produk yang sering di produksi adalah pada bulan

September cluster 1 sebanyak 8474.600 items dan cluster 2 sebanyak 25106.750 items. Dan produk yang jarang di produksi adalah pada bulan oktober cluster 1 sebanyak 7420.733 items dan cluster 2 sebanyak 2939.250 items.

2. Semakin banyak data produksi produk yang diinput, maka semakin banyak clustering centroid yang terbentuk di proses metode K-Means.
3. Menerapkan metode K-Means pada Rapidminer dilakukan dengan memasukkan data produksi bulan September dan oktober yang akan menjadi Database pada Ms.Excel, data tersebut kemudian dikoneksikan ke dalam Tools Rapidminer, dan akan diolah dan dibentuk K-Means. Setelah itu, Rapidminer akan menghasilkan produk mana yang sering (banyak) di produksi dan yang jarang (sedikit) di produksi.

V.REFERENSI

- Azhami, I., & Fauziah, R. (2020). Penerapan Rapidminer Pada Data Mining Klustering (Kasus: Distribusi Persentase Rumah Tangga Menurut Kabupaten/Kota Dan Bahan Bakar Untuk Memasak). *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, 1(2), 52-58.
- Sari, V. N., Yupianti, Y., & Maharani, D. (2018). Penerapan Metode K-Means Clustering Dalam Menentukan Predikat Kelulusan Mahasiswa Untuk Menganalisa Kualitas Lulusan. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 4(2), 133-140.
- Setiawan, R. (2017). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus: Politeknik Lp3i Jakarta). *Jurnal Lentera Ict*, 3(1), 76-92.
- Setiawan, R. (2017). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus: Politeknik Lp3i Jakarta). *Jurnal Lentera Ict*, 3(1), 76-92.
- Saragih, S. N., Safii, M., & Suhendro, D. (2021). Implementasi Metode K-Means pada Hasil Produksi Daging Jenis Ternak. *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 6(1), 235-243.
- Win, Y. M., & Oo, S. A. N. (2019). Wireless Student Attendance System using Fingerprint Sensor.
- SHAFIEL, M. F. A. B. FINGERPRINT AUTHENTICATION SYSTEM FOR EXAMINATION ATTENDANCE.
- Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 192-198.
- Sani, A. (2018). Penerapan metode k-means clustering pada perusahaan. *Jurnal Ilmiah Teknologi Informasi*, 353, 1-7.
- REGINA, S. ALGORITMA K-MEANS UNTUK CLUSTERING KUALITAS KINERJA KARYAWAN PADA PT CLARIANT ADSORBENTS INDONESIA.
- Mardalius, M. (2018). Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means.
- Sari, V. N., Yupianti, Y., & Maharani, D. (2018). Penerapan Metode K-Means Clustering Dalam Menentukan Predikat Kelulusan Mahasiswa Untuk Menganalisa Kualitas Lulusan. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 4(2), 133-140.
- Azhami, I., & Fauziah, R. (2020). Penerapan Rapidminer Pada Data Mining Klustering (Kasus: Distribusi Persentase Rumah Tangga Menurut Kabupaten/Kota Dan Bahan Bakar Untuk Memasak). *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, 1(2), 52-58.
- Sari, R. W., Wanto, A., & Windarto, A. P. (2018). Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi). *KOMIK*

- (Konferensi Nasional Teknologi Informasi dan Komputer), 2(1).
- Hendajani, F. (2019). Perbandingan Penggunaan Aplikasi Rapid Miner Dengan Weka Untuk Penerapan Metode K-Means Clustering Pada Pengelompokan Penderita Demensia. *Jurnal Ilmiah KOMPUTASI*, 18(2), 149-162.
- Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi. *CESS (Journal of Computer Engineering, System and Science)*, 3(2), 173-178.
- Siregar, A. M., Kom, S., Puspabhuana, M. K. D. A., Kom, S., & Kom, M. (2017). Data Mining: Pengolahan Data Menjadi Informasi dengan RapidMiner. CV Kekata Group.
- Arifin, J. (2019). Microsoft Excel untuk Menyusun Laporan Keuangan. Elex Media Komputindo.
- Aryati, S. IMPLEMENTASI KONSEP OBJECT ORIENTED PROGRAMMING PADA SISTEM LAUNDRY MENGGUNAKAN NETBEANS IDE 8.2.
- Sitepu, R., Irmeilyana, I., & Gultom, B. (2011). Analisis cluster terhadap tingkat pencemaran udara pada sektor industri di Sumatera Selatan. *Jurnal Penelitian Sains*, 14(3).
- Indriyani, F., & Irfiani, E. (2019). Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means. *JUITA: Jurnal Informatika*, 7(2), 109-113.
- Rohilla, V., Chakraborty, S., & Singh, M. S. (2019, October). Data Clustering using Bisecting K-Means. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 80-83). IEEE.
- Liu, J., Wang, D., Ma, Z., & Xu, B. (2019, December). Improved K-means clustering algorithm for screw locking classification. In *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (Vol. 1, pp. 1955-1958). IEEE.
- Mardalius, M. (2018). Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means.
- Normah, N., Nurajizah, S., & Salbinda, A. (2021). Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten. *Jurnal Teknik Komputer*, 7(2), 158-163.
- Sari, V. N., Yupianti, Y., & Maharani, D. (2018). Penerapan Metode K-Means Clustering Dalam Menentukan Predikat Kelulusan Mahasiswa Untuk Menganalisa Kualitas Lulusan. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 4(2), 133-140.
- Sani, A. (2018). Penerapan metode k-means clustering pada perusahaan. *Jurnal Ilmiah Teknologi Informasi*, 353, 1-7.
- Mudzakkir, B. D. (2018). Pengelompokan Data Penjualan Produk Pada PT Adv
- Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Asessment Center untuk Clustering Program SDP. *CESS (Journal of Computer Engineering, System and Science)*, 3(1), 87-93.
- Sulistiyarini, D. H., Novareza, O., & Darmawan, Z. (2018). Pengantar Proses Manufaktur untuk Teknik Industri. Universitas Brawijaya Press.
- Pramana, S., Yuniarto, B., Mariyah, S., Santoso, I., & Nooraeni, R. (2018). Data Mining dengan R Konsep Serta Implementasi. Jakarta: InMedia.
- Nofriansyah, D., & Nurcahyo, G. W. (2015). Algoritma Data Mining dan Pengujian. Deepublish.
- Prasetyowati, E. (2017). DATA MINING Pengelompokan Data untuk Informasi dan Evaluasi (Vol. 29). Duta Media Publishing.

- Buulolo, E. (2020). *Data Mining Untuk Perguruan Tinggi*. Deepublish.
- Arifin, J. (2019). *Microsoft Excel untuk Menyusun Laporan Keuangan*. Elex Media Komputindo.
- Suntoro, J. (2019). *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP*. Elex Media Komputindo.
- Arhami, M., Kom, M., & Muhammad Nasir, S. T. (2020). *Data Mining-Algoritma dan Implementasi*. Penerbit Andi.
- Hofmann, M., & Klinkenberg, R. (Eds.). (2016). *RapidMiner: Data mining use cases and business analytics applications*. CRC Press.